Neural Signatures of Dynamic Trust in AVs

Yinsu Zhang[1], Anthony D. McDonald[1], Ranjana K. Mehta[1]
[1]Department of Industrial & Systems Engineering, Texas A&M University, TX, USA

TEXAS A&M UNIVERSITY
Wm Michael Barnes '64 Department of Industrial & Systems Engineering

NeuroErgonomics Laboratory

Objectives:

To investigate and understand the effects of trust dynamics during AV driving tasks using objective measurement (brain activities), in order to achieve calibrated trust between driver & AV, and to avoid abuse or disuse of such technology.

Background:

Many automated vehicles (AVs) on road are equipped with SAE level 2 automation (SAE International 2021), requiring constant supervision by human driver. Drivers tend to over-trust and abuse the autonomy and do not provide enough supervision (Parasuraman and Manzey 2010); meanwhile, 86% of drivers reported that they are afraid about riding in an AV (Edmonds 2021), suggesting that under-trusting is also an issue. In this study, we investigated the effects of trust dynamics during AV driving tasks on brain activity, measured objectively using functional near-infrared spectroscopy (fNIRS), alongside subjective trust ratings, the current state-of-the-art for trust measure.

Trust Build 1

Trust Build 1

Trust Breach 1

Trust Breach 2

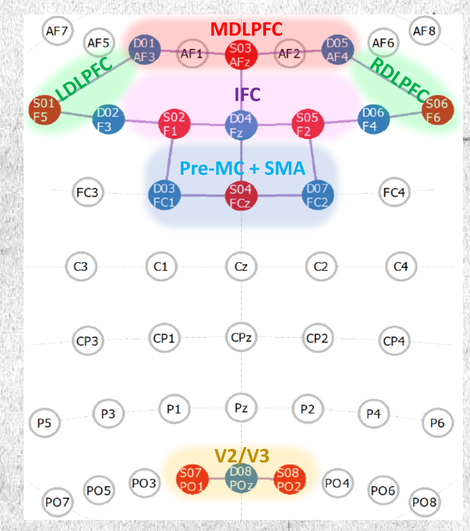
Trust Repair

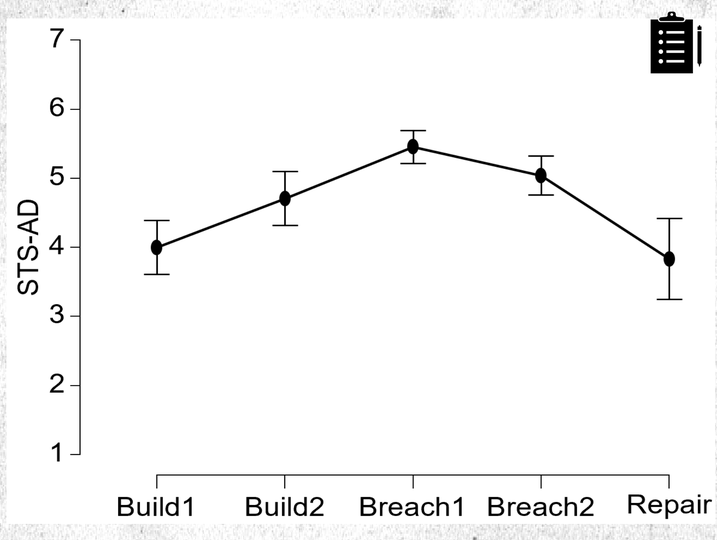
Methods:

- 60 participants, 25-50 years old, **sex balanced** (On-Going)
- An immersive driving simulator (Realtime Technologies, USA)
- 70-min experimental drive contains **five scenarios**
- Subjective trust survey: Situational Trust Scale for Automated Driving (STS-AD)
- Functional near-infrared spectroscopy (fNIRS) for **neural signatures**

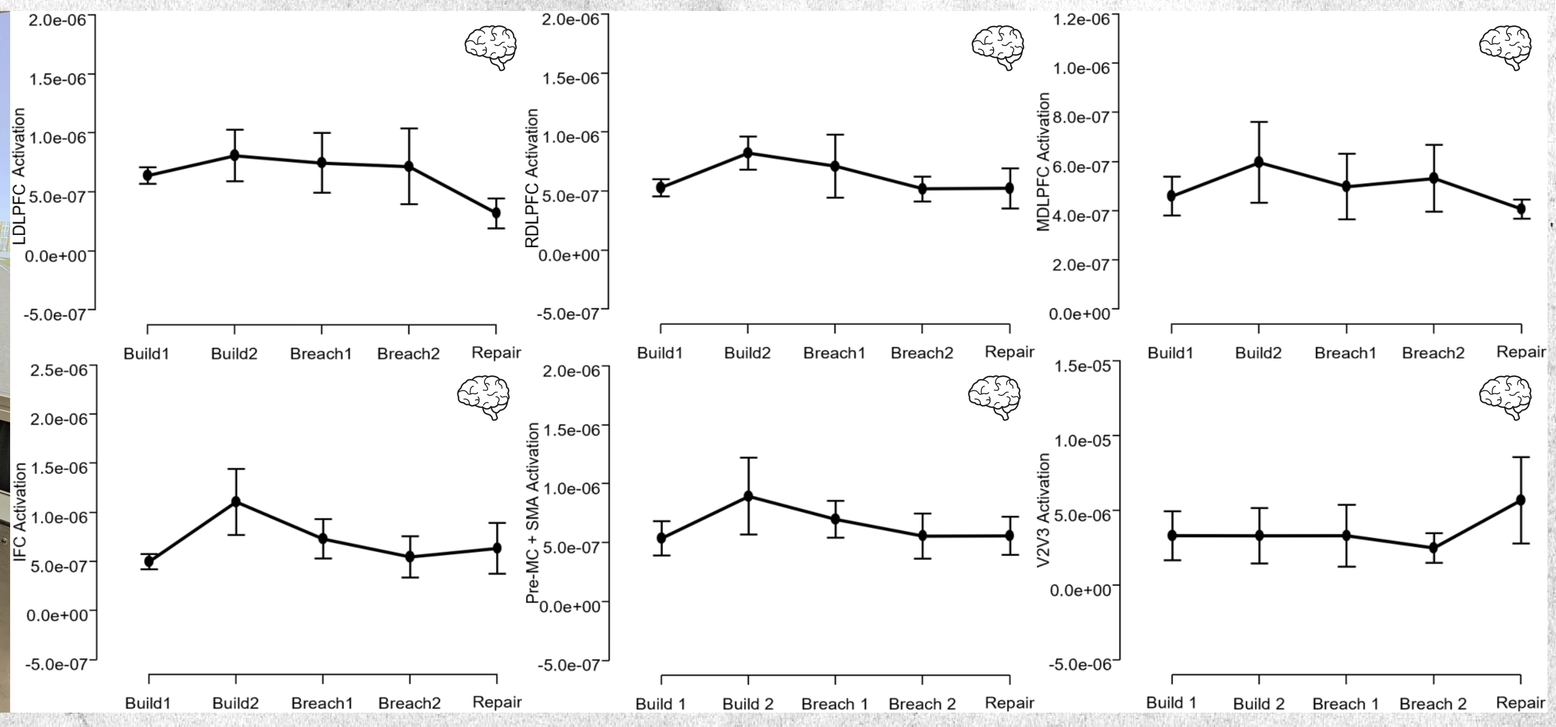
Metrics:

- Self-reported trust (STS-AD) **survey** after each scenario
- **fNIRS peak cortical activation** during each scenario
- **fNIRS functional connectivity** of 5-min intervals after each scenario



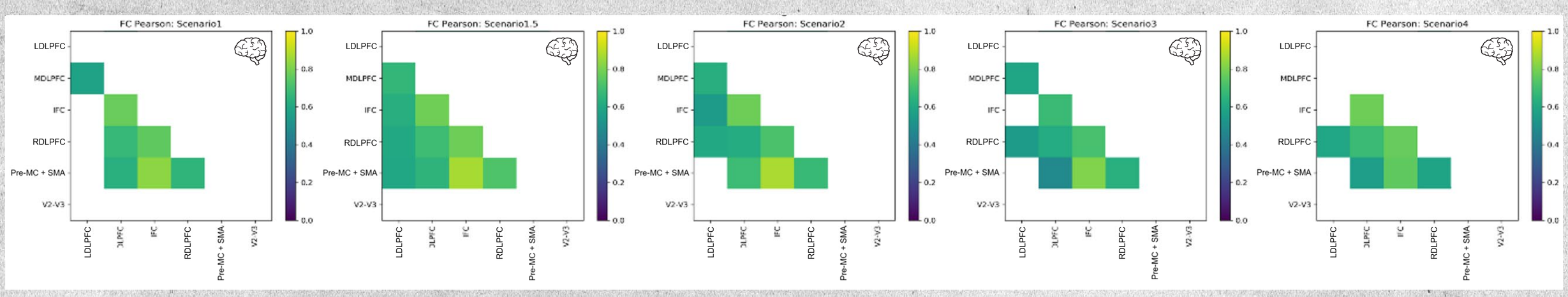


Preliminary Results:



Subjective trust results provided insights for the effects of different traffic scenarios on AV trust, demonstrating that different people can have different expectations for the AV abilities; trust repair becomes hard after major trust violations, and people tend to take over when they observe critical emergency situations.

fNIRS peak activation analysis findings are in line with previous human-robot collaboration studies (Hopko and Mehta 2021); regions involving **motor planning**, **attention** and **work memories** were potentially related to trust dynamics, and the increase in **visual information processing** can be explained by the novel stimuli introduced during the jaywalking scenario.



fNIRS functional connectivity analysis revealed the **LDLPFC** area as the predominant region that gained and lost connections during trust increase and decrease, respectively.

References:

Edmonds, Ellen. 2021. "AAA: Today's Vehicle Technology Must Walk So Self-Driving Cars Can Run." AAA Newsroom. https://newsroom.aaa.com/2021/02/aaa-todays-vehicle-technology-must-walk-so-self-driving-cars-can-run/.

Hopko, Sarah K., and Ranjana K. Mehta. 2021. "Neural Correlates of Trust in Automation: Considerations and Generalizability Between Technology Domains." Frontiers in Neuroergonomics 2 (September): 1–10.

Parasuraman, Raja, and Dietrich H. Manzey. 2010. "Complacency and Bias in Human Use of Automation: An Attentional Integration." Human Factors 52 (3): 381–410.

SAE International. 2021. "J3016 - Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles." SAE International.

Grant #: 2035367